

Extraction and Evaluation of Medication Data from Electronic Dental Records

Yue Wang^a, Zasim Siddiqui^a, Anand Krishnan^{a,b}, Jay Patel^{a,b}, Thankam Thyvalikakath^{a,c}

^aDental Informatics Core, Indiana University School of Dentistry, Indianapolis, Indiana, USA

^bIndiana University School of Informatics and Computing, Indianapolis, Indiana, USA

^cRegenstrief Institute, Indianapolis, Indiana, USA

Abstract

With an increase in the geriatric population, dental care professionals are presented with older patients who are managing their comorbidities using multiple medications. In this study, we developed a system to extract medication information from electronic dental records (EDRs) and provided patient distribution by the number of medications.

Keywords:

Dental Informatics, Electronic Dental Record, Natural Language Processing

Introduction

Preoperative evaluation of dental patients' medication history is a vital component that can significantly affect the dental care provided. Medication information extraction using natural language processing (NLP) was the theme for the 2009 Informatics for Integrating Biology and the Bedside (i2b2) challenge[1]. Following in the footsteps of the i2b2 challenge, we evaluated the medication lexical variants that existed in our dataset and the patient distribution according to the number of medications.

Methods

Cohort Identification

Electronic dental records (EDRs) were extracted from Indiana University School of Dentistry (IUSD) for adult patients who went under comprehensive oral evaluation from January 1st, 2009 to December 31st, 2011. A total of 11,220 unique patients de-identified records were used for this study.

Medication-related Term Identification and Extraction

Three data fields from medical history form was used for this study: "Do you take any medications?", "What medications are you currently taking?" and "Have you taken any other medication in the past 5 years". The data were first tokenized to generate singular terms and any duplications of these terms were removed. Two reviewers independently identified the medication related term (terms containing drugs name: either brand name or generic name) and non-medication related term (any term except medications such as noun (patients), verbs (taking), etc.). Any disagreement between them were resolved through discussion. These medication-related terms was then ran against the 11,220 patient records.

Descriptive statistics were used for this study. Medication lexicon profile and patient distribution are presented.

Results

Table 1. Medication Lexicon Profile Generated by Reviewers
Medication names were an exact match to a drug name. Medication-related terms were lexical variants.

	N	Percentage
Medication names	1,022	18.35%
Medication-related Terms	4,547	81.65%
Total	5,569	100.00%

Table 2. Patient Distribution by Number of Medications.

Number of Medications	Male	Female	Total
0	2,725	2,564	5,289
1 to 3	1,456	1,854	3,310
4 to 6	531	825	1,356
7 to 9	262	284	546
10 to 12	126	197	323
13 to 15	62	97	159
16+	50	106	156
Total	5,212	5,927	11,220

Conclusion

This study demonstrated the need to capture lexical variants for medication extraction from EDRs. Additionally, this study showed that majority of dental patients are taking at least one medication.

Acknowledgements

We would like to thank Dr. Mei Song for her valuable comments on the manuscript.

References

- [1] O. Uzuner, I. Solti, and E. Cadag, Extracting medication information from clinical text, *J Am Med Inform Assoc* 17 (2010), 514-518.

Address for Correspondence

Thankam Paul Thyvalikakath, DMD, MDS, PhD
Associate Professor & Director of Dental Informatics Core
Research Scientist, Regenstrief Institute, Inc.
Indiana University School of Dentistry, IUPUI
1121 W Michigan Street, Room DS314
Indianapolis, IN 46202
Email: tpt@iu.edu Phone: (317) 274-5460